

Klasifikasi Data Kanker Hati Menggunakan Metode Improved Random Forest-based Rule Extraction = Liver Cancer Classification Using Improved Random Forest-Based Rule Extraction

Vabiyana Safira Desdhanty, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20529168&lokasi=lokal>

Abstrak

Kanker adalah salah satu penyebab kematian utama di dunia, dengan jumlah kematian sekitar sepuluh juta kematian setiap tahun. Kanker hati menempati peringkat keenam untuk jenis kanker yang umum terjadi pada pria dan wanita. Menurut penelitian, pendeteksian dini penting untuk mencegah penyebaran kanker ke organ lain. Hal ini menyebabkan penggunaan machine learning di bidang medis untuk mengklasifikasikan data kanker agar menghasilkan diagnosis yang tepat. Namun ada kalanya dibutuhkan lebih dari satu algoritma untuk meningkatkan akurasi. Maka dari itu, penelitian ini bertujuan untuk menganalisis pengaruh Genetic Algorithm sebagai penyetelan hyperparameter untuk nilai akurasi, Penggunaan Random Forest dengan Genetic Algorithm sebagai penyetel hyperparameter memberikan akurasi sebesar 85% dengan data testing 90%. Sementara untuk Random Forest saja, hasil akurasi tertinggi adalah 73% dengan data testing sebesar 40%.

.....Cancer is one of the leading causes of mortality worldwide, with approximately ten million deaths each year. Liver cancer is the sixth most common type that occurs in both men and women. According to scientific studies, early detection is important to prevent the spread of this ailment to other organs. This led to Machine Learning in medical fields for classifying cancer data to produce an accurate diagnosis. However, there are times where just one machine learning algorithm is not giving a good accuracy score. Therefore, this study aims to analyze the effect of using Genetic Algorithm as hyperparameter tuning in terms of the accuracy level. The usage of Random Forest with Genetic Algorithm as the hyperparameter tuning algorithm gives the accuracy of 85% with 90% data testing. Meanwhile, with Random Forest alone, the highest accuracy score is 73% with 40% testing data.